



SEQUENCE LISTING

<110> D'HULST, CHRISTOPHE
BALL, STEVEN

<120> STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF
INTEREST, A METHOD OF OBTAINING THEM, AND THEIR USES

<130> 410.020

<140> 09/980,771

<141> 2002-01-10

<150> PCT/FR00/01384

<151> 2000-05-19

<150> FR 99/06494

<151> 1999-05-21

<160> 15

<170> PatentIn Ver. 3.2

<210> 1

<211> 3117

<212> DNA

<213> Chlamydomonas reinhardtii

<400> 1

```
gccaggcgga gcgtatggct gttgcctcta ccagccgccc cagcagcgcg cgtcctatcg 60
tcatcaatgc cgcgtcgttc ggtgtcaaga agaccgcgaa ccagctgctg cgtgagcttg 120
ctcgtggctc cgcacgcaag tccacctcgc gctcggtgtg tactggtgcc actggtgcca 180
cttgcgcgct ggacatcgct atggttgctg ctgaggctgc cccttggtcc aagacgggcg 240
gcctggggca tgtgactggt ggcctgccta ttgagctggt caagcgcggc caccgcgtca 300
tgaccattgc ccctcgctac gaccagtagc ctgacgcctg ggacacctcg gtggtcgtgg 360
acatcatggg cgagaaggct cgctacttcc actccatcaa gaagggcgctg caccgcgtgt 420
ggattgacca cccctggttc ctggccaagg tctggggcaa gaccggctcc aagctgtacg 480
gcccccgctc cggcgctgac tacctggaca accacaagcg cttcgccctg ttctgcaagg 540
ccgctattga ggctgcccgc gtgctgccct tcggccccgg cgaggactgc gtcttcgtgg 600
ccaacgactg gcaactccgc ctggtgcccg tcctgctgaa ggacgagtag cagcccaagg 660
gccagttcac caaggccaag tcggtgctgg ctatccacaa catcgcttc cagggccgca 720
tgtgggagga ggctttcaag gacacgaagc tgccccaggc cgcctttgac aagctggcct 780
tctcgagcgg ctatgccaa gttttacttg aggccacccc catggaggag gacgagaagc 840
ccccgctgac gggaaagacc tacaagaaga tcaactggct gaagggtggc attatcgccg 900
ccgacaagct ggtgactgtg tcgcccact acgcgaccga gatcgctgcc gatgccgccg 960
gcggtgtgga gctggacacc gtcacccgcg ccaaggcat tgagggcatt gtgaacggca 1020
tggacattga ggagtggaa cccaagaccg acaagttcct gtctgcgccc tacgaccaga 1080
acagcgtcta cgccgggcaag gccgcccga aggaggccct gcaggccgag ctgggcctgc 1140
ctgtggaccc caccgcccc ctgttcgcct tcatcgggcg cctggaggag cagaagggtg 1200
tgacatcat cctggccgcc ctgcccaga tcctggccac cccaaggtg cagatcgcca 1260
tcctgggtac cggcaaggcc gcctacgaga agctggtgaa cgccatcggc accaagtaca 1320
agggccgcgc caaggcgctg gtcaagttct cggcgccctt ggcgcacatg ctcaccgccg 1380
gcgccgactt catgctgggt ccctcgcgct tcgagccctg cggcctgatc cagctgcacg 1440
ccatgcacta cggtagcgtg cccgtggtag cctccaccgg cggcctggtc gacaccgtca 1500
aggagggcgt caccggcttc cacatgggcg ccctgaaccc cgacaagctg gacgaggctg 1560
acgccgacgc cctggccgcc accgtgcgcc gtgccagcga ggtgtttgcg ggcggccgct 1620
accccagat ggtggccaac tgcacagcc aggaacctgt ctggtccaag cccgcccaga 1680
agtgggaggg cctgctggag gaggtggtgt acggcaaggg cggcgtggcc accgccaaga 1740
```

```

aggaggagat caaggtgccc gttgccgaga agatccccgg cgacctgccc gccgtgtcct 1800
acgcccccaa caccctgaag cccgtgtccg cctccgtgga gggcaacggc gccgcccgcgc 1860
ccaaggtcgg caccaccgcc cccgccatgg gcgcgtggcg cgcgaccacc ccctcggggcc 1920
cctcgcccgc cgccgccacc cccaaggtga ccacctacaa gcccgccttg cccgccaccg 1980
ccaagcccaa gaccgctggc ctcaagctgg ccggtgaggc ctccaccacc tcgacctcgg 2040
agaacggcgc tgcctccaac ggcaacggca acggtgcctc ggctccaag acctcggctg 2100
ccaagcccct ggtctccgcc gccaccgca agtccgccta aagcggcagt agccgcagag 2160
gcggcgacag catgagcggc tcgaccaaag ctgtggcagg aacggctgta gcagcggcag 2220
gcggccgcca ccggcgagga gcaggcttgc ggcagcgagg gcgatgagct tagcggccgt 2280
gagcatggca ggcggaacg tgtgtactga aatgtggtgc atgagagtgt cgtgctgtaa 2340
tgaagtcggt tttgcgagac cggagaaacg ccggtttggt tttgtagtgc agggcctgtg 2400
gtttcggttt tgcccaagtc caaaagaaga gtaacgaaac tgtagcagta gcagagcact 2460
tgcgcggcgc ggcgaccacg ccggcccgtg cgcagcctgt cctgccctca gccttgtgat 2520
tcggcggcaa gagggcgggt ctgtacactc catccattcc aggatTTTTg caggctgcct 2580
gagagtttgc cattttgtgg gacgtgagcg gcgggacggc cgcgccgggc tctcctaccg 2640
cctccggcaa cggagaagtg ggaggcgtg tagcccggtg accccccaat gtagaggatg 2700
ggatacataa gagcgtgtgg aatggtggtg aaagaggagg ggcctgggtc gccctcgat 2760
ggttttgttg aggtgcagac ggcaccgtcg gcgtcaaagg ccctcgcaag gcccggtgc 2820
cttgggctca tttttggtgc ccgtcgatga tgagagattg gccagcggtt ttttgaggct 2880
ggctcgaagc gaggtttgt ggaagtggag cgaggagggt tggagaaaga ggcggacatg 2940
cttgactgga ggtacacaaa gtggagcgtg cgacggcacg gaggcattgg cggactattg 3000
accagtagt gtggaagta gttggacctg aattcTTTga ggtaccgcg cattaatccg 3060
tgagagagta acaaagatgg cacctgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 3117

```

<210> 2

<211> 2124

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Fragment of the complete sequence of cDNA coding for the GBSSI of *Chlamydomonas reinhardtii*

<220>

<221> CDS

<222> (1)..(2124)

<400> 2

```

atg gct gtt gcc tct acc agc cgc ccc agc agc gcg cgt cct atc gtc 48
Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val
  1             5             10             15

atc aat gcc gcg tcg ttc ggt gtc aag aag acc gcg aac cag ctg ctg 96
Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu
      20             25             30

cgt gag ctt gct cgt ggc tcc gca cgc aag tcc acc tcg cgc tcg gct 144
Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala
      35             40             45

gtt act ggt gcc act ggt gcc act tgc gcg ctg gac atc gtg atg gtt 192
Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val
      50             55             60

gct gct gag gtc gcc cct tgg tcc aag acg ggc ggc ctg ggc gat gtg 240
Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val

```

65	70	75	80	
act ggt ggc ctg cct att gag ctg gtc aag cgc ggc cac cgc gtc atg	288			
Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met				
85 90 95				
acc att gcc cct cgc tac gac cag tac gct gac gcc tgg gac acc tcg	336			
Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser				
100 105 110				
gtg gtc gtg gac atc atg ggc gag aag gtc cgc tac ttc cac tcc atc	384			
Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile				
115 120 125				
aag aag ggc gtg cac cgc gtg tgg att gac cac ccc tgg ttc ctg gcc	432			
Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala				
130 135 140				
aag gtc tgg ggc aag acc ggc tcc aag ctg tac ggc ccc cgc tcc ggc	480			
Lys Val Trp Gly Lys Thr Gly Ser Lys Leu Tyr Gly Pro Arg Ser Gly				
145 150 155 160				
gct gac tac ctg gac aac cac aag cgc ttc gcc ctg ttc tgc aag gcc	528			
Ala Asp Tyr Leu Asp Asn His Lys Arg Phe Ala Leu Phe Cys Lys Ala				
165 170 175				
gct att gag gct gcc cgc gtg ctg ccc ttc ggc ccc ggc gag gac tgc	576			
Ala Ile Glu Ala Ala Arg Val Leu Pro Phe Gly Pro Gly Glu Asp Cys				
180 185 190				
gtc ttc gtg gcc aac gac tgg cac tcc gcc ctg gtg ccc gtc ctg ctg	624			
Val Phe Val Ala Asn Asp Trp His Ser Ala Leu Val Pro Val Leu Leu				
195 200 205				
aag gac gag tac cag ccc aag ggc cag ttc acc aag gcc aag tcg gtg	672			
Lys Asp Glu Tyr Gln Pro Lys Gly Gln Phe Thr Lys Ala Lys Ser Val				
210 215 220				
ctg gct atc cac aac atc gcc ttc cag ggc cgc atg tgg gag gag gct	720			
Leu Ala Ile His Asn Ile Ala Phe Gln Gly Arg Met Trp Glu Glu Ala				
225 230 235 240				
ttc aag gac acg aag ctg ccc cag gcc gcc ttt gac aag ctg gcc ttc	768			
Phe Lys Asp Thr Lys Leu Pro Gln Ala Ala Phe Asp Lys Leu Ala Phe				
245 250 255				
tcg gac ggc tat gcc aag gtt tac act gag gcc acc ccc atg gag gag	816			
Ser Asp Gly Tyr Ala Lys Val Tyr Thr Glu Ala Thr Pro Met Glu Glu				
260 265 270				
gac gag aag ccc ccg ctg acg gga aag acc tac aag aag atc aac tgg	864			
Asp Glu Lys Pro Pro Leu Thr Gly Lys Thr Tyr Lys Lys Ile Asn Trp				
275 280 285				
ctg aag ggt ggc att atc gcc gcc gac aag ctg gtg act gtg tcg ccc	912			
Leu Lys Gly Gly Ile Ile Ala Ala Asp Lys Leu Val Thr Val Ser Pro				
290 295 300				

aac tac gcg acc gag atc gct gcc gat gcc gcc gcc ggt gtg gag ctg	960
Asn Tyr Ala Thr Glu Ile Ala Ala Asp Ala Ala Gly Gly Val Glu Leu	
305 310 315 320	
gac acc gtc atc cgc gcc aag ggc att gag ggc att gtg aac ggc atg	1008
Asp Thr Val Ile Arg Ala Lys Gly Ile Glu Gly Ile Val Asn Gly Met	
325 330 335	
gac att gag gag tgg aac ccc aag acc gac aag ttc ctg tct gcg ccc	1056
Asp Ile Glu Glu Trp Asn Pro Lys Thr Asp Lys Phe Leu Ser Ala Pro	
340 345 350	
tac gac cag aac agc gtc tac gcc ggc aag gcc gcc gcc aag gag gcc	1104
Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala	
355 360 365	
ctg cag gcc gag ctg ggc ctg cct gtg gac ccc acc gcc ccc ctg ttc	1152
Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe	
370 375 380	
gcc ttc atc ggc cgc ctg gag gag cag aag ggt gtg gac atc atc ctg	1200
Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu	
385 390 395 400	
gcc gcc ctg ccc aag atc ctg gcc acc ccc aag gtg cag atc gcc atc	1248
Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile	
405 410 415	
ctg ggt acc ggc aag gcc gcc tac gag aag ctg gtg aac gcc atc ggc	1296
Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly	
420 425 430	
acc aag tac aag ggc cgc gcc aag ggc gtg gtc aag ttc tcg gcg ccc	1344
Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro	
435 440 445	
ctg gcg cac atg ctc acc gcc ggc gcc gac ttc atg ctg gtg ccc tcg	1392
Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser	
450 455 460	
cgc ttc gag ccc tgc ggc ctg atc cag ctg cac gcc atg cac tac ggt	1440
Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly	
465 470 475 480	
acc gtg ccc gtg gta gcc tcc acc ggc ggc ctg gtc gac acc gtc aag	1488
Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys	
485 490 495	
gag ggc gtc acc ggc ttc cac atg ggc gcc ctg aac ccc gac aag ctg	1536
Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu	
500 505 510	
gac gag gct gac gcc gac gcc ctg gcc gcc acc gtg cgc cgt gcc agc	1584
Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser	
515 520 525	
gag gtg ttt gcg ggc ggc cgc tac ccc gag atg gtg gcc aac tgc atc	1632
Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile	

530	535	540	
agc cag gac ctg tcc tgg tcc aag ccc gcc cag aag tgg gag ggc ctg			1680
Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu			
545	550	555	560
ctg gag gag gtg gtg tac ggc aag ggc ggc gtg gcc acc gcc aag aag			1728
Leu Glu Glu Val Val Tyr Gly Lys Gly Val Ala Thr Ala Lys Lys			
	565	570	575
gag gag atc aag gtg ccc gtt gcc gag aag atc ccc ggc gac ctg ccc			1776
Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro			
	580	585	590
gcc gtg tcc tac gcc ccc aac acc ctg aag ccc gtg tcc gcc tcc gtg			1824
Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val			
	595	600	605
gag ggc aac ggc gcc gcc gcg ccc aag gtc ggc acc acc gcc ccc gcc			1872
Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala			
	610	615	620
atg ggc gcg tgg cgc gcg acc acc ccc tcg ggc ccc tcg ccc gcc gcc			1920
Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala			
	625	630	635
gcc acc ccc aag gtg acc acc tac aag ccc gcc ctg ccc gcc acc gcc			1968
Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala			
	645	650	655
aag ccc aag acc gct ggc ctc aag ctg gcc ggt gag gcc tcc acc acc			2016
Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr			
	660	665	670
tcg acc tcg gag aac ggc gct gcc tcc aac ggc aac ggc aac ggt gcc			2064
Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala			
	675	680	685
tcg gcc tcc aag acc tcg gct gcc aag ccc ctg gtc tcc gcc gcc acc			2112
Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr			
	690	695	700
cgc aag tcc gcc			2124
Arg Lys Ser Ala			
705			

<210> 3

<211> 708

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Amino acid sequence
of the GBSSI of Chlamydomonas reinhardtii

<400> 3

Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val
1 5 10 15

Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu
20 25 30

Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala
35 40 45

Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val
50 55 60

Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val
65 70 75 80

Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met
85 90 95

Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser
100 105 110

Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile
115 120 125

Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala
130 135 140

Lys Val Trp Gly Lys Thr Gly Ser Lys Leu Tyr Gly Pro Arg Ser Gly
145 150 155 160

Ala Asp Tyr Leu Asp Asn His Lys Arg Phe Ala Leu Phe Cys Lys Ala
165 170 175

Ala Ile Glu Ala Ala Arg Val Leu Pro Phe Gly Pro Gly Glu Asp Cys
180 185 190

Val Phe Val Ala Asn Asp Trp His Ser Ala Leu Val Pro Val Leu Leu
195 200 205

Lys Asp Glu Tyr Gln Pro Lys Gly Gln Phe Thr Lys Ala Lys Ser Val
210 215 220

Leu Ala Ile His Asn Ile Ala Phe Gln Gly Arg Met Trp Glu Glu Ala
225 230 235 240

Phe Lys Asp Thr Lys Leu Pro Gln Ala Ala Phe Asp Lys Leu Ala Phe
245 250 255

Ser Asp Gly Tyr Ala Lys Val Tyr Thr Glu Ala Thr Pro Met Glu Glu
260 265 270

Asp Glu Lys Pro Pro Leu Thr Gly Lys Thr Tyr Lys Lys Ile Asn Trp
275 280 285

Leu Lys Gly Gly Ile Ile Ala Ala Asp Lys Leu Val Thr Val Ser Pro
290 295 300

Asn Tyr Ala Thr Glu Ile Ala Ala Asp Ala Ala Gly Gly Val Glu Leu
 305 310 315 320
 Asp Thr Val Ile Arg Ala Lys Gly Ile Glu Gly Ile Val Asn Gly Met
 325 330 335
 Asp Ile Glu Glu Trp Asn Pro Lys Thr Asp Lys Phe Leu Ser Ala Pro
 340 345 350
 Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala
 355 360 365
 Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe
 370 375 380
 Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu
 385 390 395 400
 Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile
 405 410 415
 Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly
 420 425 430
 Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro
 435 440 445
 Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser
 450 455 460
 Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly
 465 470 475 480
 Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys
 485 490 495
 Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu
 500 505 510
 Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser
 515 520 525
 Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile
 530 535 540
 Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu
 545 550 555 560
 Leu Glu Glu Val Val Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys
 565 570 575
 Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro
 580 585 590
 Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val
 595 600 605
 Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala

610	615	620
Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala		
625	630	635 640
Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala		
	645	650 655
Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr		
	660	665 670
Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala		
	675	680 685
Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr		
	690	695 700
Arg Lys Ser Ala		
705		

<210> 4
 <211> 1953
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Fragment of the complete sequence of cDNA coding for the GBSSI of *Chlamydomonas reinhardtii* and coding for the mature GBSSI protein

<220>
 <221> CDS
 <222> (1)..(1953)

<400> 4	
gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag	48
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys	
1 5 10 15	
acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc	96
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val	
20 25 30	
aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac	144
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr	
35 40 45	
gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag	192
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys	
50 55 60	
gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att	240
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile	
65 70 75 80	

gac	cac	ccc	tgg	ttc	ctg	gcc	aag	gtc	tgg	ggc	aag	acc	ggc	tcc	aag	288
Asp	His	Pro	Trp	Phe	Leu	Ala	Lys	Val	Trp	Gly	Lys	Thr	Gly	Ser	Lys	
				85					90					95		
ctg	tac	ggc	ccc	cgc	tcc	ggc	gct	gac	tac	ctg	gac	aac	cac	aag	cgc	336
Leu	Tyr	Gly	Pro	Arg	Ser	Gly	Ala	Asp	Tyr	Leu	Asp	Asn	His	Lys	Arg	
			100					105					110			
ttc	gcc	ctg	ttc	tgc	aag	gcc	gct	att	gag	gct	gcc	cgc	gtg	ctg	ccc	384
Phe	Ala	Leu	Phe	Cys	Lys	Ala	Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro	
		115					120					125				
ttc	ggc	ccc	ggc	gag	gac	tgc	gtc	ttc	gtg	gcc	aac	gac	tgg	cac	tcc	432
Phe	Gly	Pro	Gly	Glu	Asp	Cys	Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser	
	130					135					140					
gcc	ctg	gtg	ccc	gtc	ctg	ctg	aag	gac	gag	tac	cag	ccc	aag	ggc	cag	480
Ala	Leu	Val	Pro	Val	Leu	Leu	Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln	
145					150					155					160	
ttc	acc	aag	gcc	aag	tcg	gtg	ctg	gct	atc	cac	aac	atc	gcc	ttc	cag	528
Phe	Thr	Lys	Ala	Lys	Ser	Val	Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln	
				165					170					175		
ggc	cgc	atg	tgg	gag	gag	gct	ttc	aag	gac	acg	aag	ctg	ccc	cag	gcc	576
Gly	Arg	Met	Trp	Glu	Glu	Ala	Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala	
			180					185					190			
gcc	ttt	gac	aag	ctg	gcc	ttc	tcg	gac	ggc	tat	gcc	aag	gtt	tac	act	624
Ala	Phe	Asp	Lys	Leu	Ala	Phe	Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr	
		195					200					205				
gag	gcc	acc	ccc	atg	gag	gag	gac	gag	aag	ccc	ccg	ctg	acg	gga	aag	672
Glu	Ala	Thr	Pro	Met	Glu	Glu	Asp	Glu	Lys	Pro	Pro	Leu	Thr	Gly	Lys	
	210					215					220					
acc	tac	aag	aag	atc	aac	tgg	ctg	aag	ggc	att	atc	gcc	gcc	gac		720
Thr	Tyr	Lys	Lys	Ile	Asn	Trp	Leu	Lys	Gly	Gly	Ile	Ile	Ala	Ala	Asp	
225					230				235					240		
aag	ctg	gtg	act	gtg	tcg	ccc	aac	tac	gcg	acc	gag	atc	gct	gcc	gat	768
Lys	Leu	Val	Thr	Val	Ser	Pro	Asn	Tyr	Ala	Thr	Glu	Ile	Ala	Ala	Asp	
				245					250				255			
gcc	gcc	ggc	ggc	gtg	gag	ctg	gac	acc	gtc	atc	cgc	gcc	aag	ggc	att	816
Ala	Ala	Gly	Gly	Val	Glu	Leu	Asp	Thr	Val	Ile	Arg	Ala	Lys	Gly	Ile	
			260					265					270			
gag	ggc	att	gtg	aac	ggc	atg	gac	att	gag	gag	tgg	aac	ccc	aag	acc	864
Glu	Gly	Ile	Val	Asn	Gly	Met	Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr	
		275				280						285				
gac	aag	ttc	ctg	tct	gcg	ccc	tac	gac	cag	aac	agc	gtc	tac	gcc	ggc	912
Asp	Lys	Phe	Leu	Ser	Ala	Pro	Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly	
	290					295					300					
aag	gcc	gcc	gcc	aag	gag	gcc	ctg	cag	gcc	gag	ctg	ggc	ctg	cct	gtg	960
Lys	Ala	Ala	Ala	Lys	Glu	Ala	Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val	

305	310	315	320	
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag				1008
Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln	325	330	335	
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc				1056
Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr	340	345	350	
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag				1104
Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu	355	360	365	
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc				1152
Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly	370	375	380	
gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc				1200
Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala	385	390	395	400
gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag				1248
Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln	405	410	415	
ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc				1296
Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly	420	425	430	
ggc ctg gtc gac acc gtc aag gag ggc gtc acc ggc ttc cac atg ggc				1344
Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe His Met Gly	435	440	445	
gcc ctg aac ccc gac aag ctg gac gag gct gac gcc gac gcc ctg gcc				1392
Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp Ala Leu Ala	450	455	460	
gcc acc gtg cgc cgt gcc agc gag gtg ttt gcg ggc ggc cgc tac ccc				1440
Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro	465	470	475	480
gag atg gtg gcc aac tgc atc agc cag gac ctg tcc tgg tcc aag ccc				1488
Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro	485	490	495	
gcc cag aag tgg gag ggc ctg ctg gag gag gtg gtg tac ggc aag ggc				1536
Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly	500	505	510	
ggc gtg gcc acc gcc aag aag gag gag atc aag gtg ccc gtt gcc gag				1584
Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu	515	520	525	
aag atc ccc ggc gac ctg ccc gcc gtg tcc tac gcc ccc aac acc ctg				1632
Lys Ile Pro Gly Asp Leu Pro Ala Val Ser Tyr Ala Pro Asn Thr Leu	530	535	540	

aag ccc gtg tcc gcc tcc gtg gag ggc aac ggc gcc gcc gcg ccc aag	1680
Lys Pro Val Ser Ala Ser Val Glu Gly Asn Gly Ala Ala Ala Pro Lys	
545 550 555 560	
gtc ggc acc acc gcc ccc gcc atg ggc gcg tgg cgc gcg acc acc ccc	1728
Val Gly Thr Thr Ala Pro Ala Met Gly Ala Trp Arg Ala Thr Thr Pro	
565 570 575	
tcg ggc ccc tcg ccc gcc gcc gcc acc ccc aag gtg acc acc tac aag	1776
Ser Gly Pro Ser Pro Ala Ala Ala Thr Pro Lys Val Thr Thr Tyr Lys	
580 585 590	
ccc gcc ctg ccc gcc acc gcc aag ccc aag acc gct ggc ctc aag ctg	1824
Pro Ala Leu Pro Ala Thr Ala Lys Pro Lys Thr Ala Gly Leu Lys Leu	
595 600 605	
gcc ggt gag gcc tcc acc acc tcg acc tcg gag aac ggc gct gcc tcc	1872
Ala Gly Glu Ala Ser Thr Thr Ser Thr Ser Glu Asn Gly Ala Ala Ser	
610 615 620	
aac ggc aac ggc aac ggt gcc tcg gcc tcc aag acc tcg gct gcc aag	1920
Asn Gly Asn Gly Asn Gly Ala Ser Ala Ser Lys Thr Ser Ala Ala Lys	
625 630 635 640	
ccc ctg gtc tcc gcc gcc acc cgc aag tcc gcc	1953
Pro Leu Val Ser Ala Ala Thr Arg Lys Ser Ala	
645 650	

<210> 5

<211> 651

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Amino acid sequence
of the GBSSI of *Chlamydomonas reinhardtii* and the
mature GBSSI protein

<400> 5

Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys
1 5 10 15

Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val
20 25 30

Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr
35 40 45

Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys
50 55 60

Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile
65 70 75 80

Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys

85										90					95				
Leu	Tyr	Gly	Pro	Arg	Ser	Gly	Ala	Asp	Tyr	Leu	Asp	Asn	His	Lys	Arg				
			100					105					110						
Phe	Ala	Leu	Phe	Cys	Lys	Ala	Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro				
		115					120					125							
Phe	Gly	Pro	Gly	Glu	Asp	Cys	Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser				
	130					135					140								
Ala	Leu	Val	Pro	Val	Leu	Leu	Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln				
145					150					155					160				
Phe	Thr	Lys	Ala	Lys	Ser	Val	Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln				
				165					170					175					
Gly	Arg	Met	Trp	Glu	Glu	Ala	Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala				
			180					185					190						
Ala	Phe	Asp	Lys	Leu	Ala	Phe	Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr				
		195					200					205							
Glu	Ala	Thr	Pro	Met	Glu	Glu	Asp	Glu	Lys	Pro	Pro	Leu	Thr	Gly	Lys				
	210					215					220								
Thr	Tyr	Lys	Lys	Ile	Asn	Trp	Leu	Lys	Gly	Gly	Ile	Ile	Ala	Ala	Asp				
225					230					235					240				
Lys	Leu	Val	Thr	Val	Ser	Pro	Asn	Tyr	Ala	Thr	Glu	Ile	Ala	Ala	Asp				
				245					250					255					
Ala	Ala	Gly	Gly	Val	Glu	Leu	Asp	Thr	Val	Ile	Arg	Ala	Lys	Gly	Ile				
			260					265					270						
Glu	Gly	Ile	Val	Asn	Gly	Met	Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr				
	275						280					285							
Asp	Lys	Phe	Leu	Ser	Ala	Pro	Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly				
	290					295					300								
Lys	Ala	Ala	Ala	Lys	Glu	Ala	Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val				
305					310					315					320				
Asp	Pro	Thr	Ala	Pro	Leu	Phe	Ala	Phe	Ile	Gly	Arg	Leu	Glu	Glu	Gln				
				325					330					335					
Lys	Gly	Val	Asp	Ile	Ile	Leu	Ala	Ala	Leu	Pro	Lys	Ile	Leu	Ala	Thr				
			340					345					350						
Pro	Lys	Val	Gln	Ile	Ala	Ile	Leu	Gly	Thr	Gly	Lys	Ala	Ala	Tyr	Glu				
		355					360					365							
Lys	Leu	Val	Asn	Ala	Ile	Gly	Thr	Lys	Tyr	Lys	Gly	Arg	Ala	Lys	Gly				
	370					375					380								
Val	Val	Lys	Phe	Ser	Ala	Pro	Leu	Ala	His	Met	Leu	Thr	Ala	Gly	Ala				
385					390					395					400				

Asp	Phe	Met	Leu	Val	Pro	Ser	Arg	Phe	Glu	Pro	Cys	Gly	Leu	Ile	Gln	405	410	415
Leu	His	Ala	Met	His	Tyr	Gly	Thr	Val	Pro	Val	Val	Ala	Ser	Thr	Gly	420	425	430
Gly	Leu	Val	Asp	Thr	Val	Lys	Glu	Gly	Val	Thr	Gly	Phe	His	Met	Gly	435	440	445
Ala	Leu	Asn	Pro	Asp	Lys	Leu	Asp	Glu	Ala	Asp	Ala	Asp	Ala	Leu	Ala	450	455	460
Ala	Thr	Val	Arg	Arg	Ala	Ser	Glu	Val	Phe	Ala	Gly	Gly	Arg	Tyr	Pro	465	470	475
Glu	Met	Val	Ala	Asn	Cys	Ile	Ser	Gln	Asp	Leu	Ser	Trp	Ser	Lys	Pro	485	490	495
Ala	Gln	Lys	Trp	Glu	Gly	Leu	Leu	Glu	Glu	Val	Val	Tyr	Gly	Lys	Gly	500	505	510
Gly	Val	Ala	Thr	Ala	Lys	Lys	Glu	Glu	Ile	Lys	Val	Pro	Val	Ala	Glu	515	520	525
Lys	Ile	Pro	Gly	Asp	Leu	Pro	Ala	Val	Ser	Tyr	Ala	Pro	Asn	Thr	Leu	530	535	540
Lys	Pro	Val	Ser	Ala	Ser	Val	Glu	Gly	Asn	Gly	Ala	Ala	Ala	Pro	Lys	545	550	555
Val	Gly	Thr	Thr	Ala	Pro	Ala	Met	Gly	Ala	Trp	Arg	Ala	Thr	Thr	Pro	565	570	575
Ser	Gly	Pro	Ser	Pro	Ala	Ala	Ala	Thr	Pro	Lys	Val	Thr	Thr	Tyr	Lys	580	585	590
Pro	Ala	Leu	Pro	Ala	Thr	Ala	Lys	Pro	Lys	Thr	Ala	Gly	Leu	Lys	Leu	595	600	605
Ala	Gly	Glu	Ala	Ser	Thr	Thr	Ser	Thr	Ser	Glu	Asn	Gly	Ala	Ala	Ser	610	615	620
Asn	Gly	Asn	Gly	Asn	Gly	Ala	Ser	Ala	Ser	Lys	Thr	Ser	Ala	Ala	Lys	625	630	635
Pro	Leu	Val	Ser	Ala	Ala	Thr	Arg	Lys	Ser	Ala						645	650	

<210> 6

<211> 1314

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Fragment of the

complete cDNA coding for the GBSSI of Chlamydomonas
reinhardtii

<220>

<221> CDS

<222> (1)..(1314)

<400> 6

gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag	48
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys	
1 5 10 15	
acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc	96
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val	
20 25 30	
aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac	144
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr	
35 40 45	
gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag	192
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys	
50 55 60	
gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att	240
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile	
65 70 75 80	
gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag	288
Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys	
85 90 95	
ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc	336
Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg	
100 105 110	
ttc gcc ctg ttc tgc aag gcc gct att gag gct gcc cgc gtg ctg ccc	384
Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro	
115 120 125	
ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc	432
Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser	
130 135 140	
gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag	480
Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln	
145 150 155 160	
ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag	528
Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln	
165 170 175	
ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc	576
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala	
180 185 190	
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act	624
Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr	

195	200	205	
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys 210 215 220			672
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp 225 230 235 240			720
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp 245 250 255			768
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile 260 265 270			816
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr 275 280 285			864
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly 290 295 300			912
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val 305 310 315 320			960
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln 325 330 335			1008
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr 340 345 350			1056
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu 355 360 365			1104
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly 370 375 380			1152
gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala 385 390 395 400			1200
gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln 405 410 415			1248
ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly 420 425 430			1296

ggc ctg gtc gac acc gtc
Gly Leu Val Asp Thr Val
435

1314

<210> 7
<211> 438
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Amino acid sequence
of the GBSSI of Chlamydomonas reinhardtii

<400> 7
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys
1 5 10 15
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val
20 25 30
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr
35 40 45
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys
50 55 60
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile
65 70 75 80
Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys
85 90 95
Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg
100 105 110
Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro
115 120 125
Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser
130 135 140
Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln
145 150 155 160
Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln
165 170 175
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala
180 185 190
Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr
195 200 205
Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys
210 215 220

Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp
 225 230 235 240
 Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp
 245 250 255
 Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile
 260 265 270
 Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr
 275 280 285
 Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly
 290 295 300
 Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val
 305 310 315 320
 Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln
 325 330 335
 Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr
 340 345 350
 Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu
 355 360 365
 Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly
 370 375 380
 Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala
 385 390 395 400
 Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln
 405 410 415
 Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly
 420 425 430
 Gly Leu Val Asp Thr Val
 435

<210> 8

<211> 1593

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Fragment of the complete cDNA coding for the GBSSI of *Chlamydomonas reinhardtii*

<220>

<221> CDS

<222> (1)..(1593)

<400> 8

gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag	48
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys	
1 5 10 15	
acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc	96
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val	
20 25 30	
aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac	144
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr	
35 40 45	
gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag	192
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys	
50 55 60	
gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att	240
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile	
65 70 75 80	
gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag	288
Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys	
85 90 95	
ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc	336
Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg	
100 105 110	
ttc gcc ctg ttc tgc aag gcc gct att gag gct gcc cgc gtg ctg ccc	384
Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro	
115 120 125	
ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc	432
Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser	
130 135 140	
gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag	480
Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln	
145 150 155 160	
ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag	528
Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln	
165 170 175	
ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc	576
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala	
180 185 190	
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act	624
Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr	
195 200 205	
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag	672
Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys	
210 215 220	

acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac	720
Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp	
225 230 235 240	
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat	768
Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp	
245 250 255	
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att	816
Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile	
260 265 270	
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc	864
Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr	
275 280 285	
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc	912
Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly	
290 295 300	
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg	960
Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val	
305 310 315 320	
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag	1008
Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln	
325 330 335	
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc	1056
Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr	
340 345 350	
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag	1104
Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu	
355 360 365	
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc	1152
Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly	
370 375 380	
gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc	1200
Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala	
385 390 395 400	
gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag	1248
Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln	
405 410 415	
ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc	1296
Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly	
420 425 430	
ggc ctg gtc gac acc gtc aag gag ggc gtc acc ggc ttc cac atg ggc	1344
Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe His Met Gly	
435 440 445	
gcc ctg aac ccc gac aag ctg gac gag gct gac gcc gac gcc ctg gcc	1392
Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp Ala Leu Ala	

450	455	460	
gcc acc gtg cgc cgt gcc agc gag gtg ttt gcg ggc ggc cgc tac ccc			1440
Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro			
465	470	475	480
gag atg gtg gcc aac tgc atc agc cag gac ctg tcc tgg tcc aag ccc			1488
Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro			
	485	490	495
gcc cag aag tgg gag ggc ctg ctg gag gag gtg gtg tac ggc aag ggc			1536
Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly			
	500	505	510
ggc gtg gcc acc gcc aag aag gag gag atc aag gtg ccc gtt gcc gag			1584
Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu			
	515	520	525
aag atc ccc			1593
Lys Ile Pro			
530			

<210> 9
 <211> 531
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Amino acid sequence
 of the GBSSI of Chlamydomonas reinhardtii

<400> 9
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys
 1 5 10 15
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val
 20 25 30
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr
 35 40 45
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys
 50 55 60
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile
 65 70 75 80
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys
 85 90 95
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg
 100 105 110
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro
 115 120 125

Phe	Gly	Pro	Gly	Glu	Asp	Cys	Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser		
130						135					140						
Ala	Leu	Val	Pro	Val	Leu	Leu	Lys	Asp	Glu	Tyr	Gln	Pro	Lys	Gly	Gln		
145					150					155					160		
Phe	Thr	Lys	Ala	Lys	Ser	Val	Leu	Ala	Ile	His	Asn	Ile	Ala	Phe	Gln		
				165					170						175		
Gly	Arg	Met	Trp	Glu	Glu	Ala	Phe	Lys	Asp	Thr	Lys	Leu	Pro	Gln	Ala		
			180					185						190			
Ala	Phe	Asp	Lys	Leu	Ala	Phe	Ser	Asp	Gly	Tyr	Ala	Lys	Val	Tyr	Thr		
		195					200					205					
Glu	Ala	Thr	Pro	Met	Glu	Glu	Asp	Glu	Lys	Pro	Pro	Leu	Thr	Gly	Lys		
	210					215					220						
Thr	Tyr	Lys	Lys	Ile	Asn	Trp	Leu	Lys	Gly	Gly	Ile	Ile	Ala	Ala	Asp		
225				230						235					240		
Lys	Leu	Val	Thr	Val	Ser	Pro	Asn	Tyr	Ala	Thr	Glu	Ile	Ala	Ala	Asp		
				245					250						255		
Ala	Ala	Gly	Gly	Val	Glu	Leu	Asp	Thr	Val	Ile	Arg	Ala	Lys	Gly	Ile		
			260					265					270				
Glu	Gly	Ile	Val	Asn	Gly	Met	Asp	Ile	Glu	Glu	Trp	Asn	Pro	Lys	Thr		
	275						280					285					
Asp	Lys	Phe	Leu	Ser	Ala	Pro	Tyr	Asp	Gln	Asn	Ser	Val	Tyr	Ala	Gly		
	290					295					300						
Lys	Ala	Ala	Ala	Lys	Glu	Ala	Leu	Gln	Ala	Glu	Leu	Gly	Leu	Pro	Val		
305					310					315					320		
Asp	Pro	Thr	Ala	Pro	Leu	Phe	Ala	Phe	Ile	Gly	Arg	Leu	Glu	Glu	Gln		
				325					330					335			
Lys	Gly	Val	Asp	Ile	Ile	Leu	Ala	Ala	Leu	Pro	Lys	Ile	Leu	Ala	Thr		
			340					345						350			
Pro	Lys	Val	Gln	Ile	Ala	Ile	Leu	Gly	Thr	Gly	Lys	Ala	Ala	Tyr	Glu		
		355					360					365					
Lys	Leu	Val	Asn	Ala	Ile	Gly	Thr	Lys	Tyr	Lys	Gly	Arg	Ala	Lys	Gly		
	370					375					380						
Val	Val	Lys	Phe	Ser	Ala	Pro	Leu	Ala	His	Met	Leu	Thr	Ala	Gly	Ala		
385					390					395					400		
Asp	Phe	Met	Leu	Val	Pro	Ser	Arg	Phe	Glu	Pro	Cys	Gly	Leu	Ile	Gln		
				405					410					415			
Leu	His	Ala	Met	His	Tyr	Gly	Thr	Val	Pro	Val	Val	Ala	Ser	Thr	Gly		
			420					425					430				

Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe His Met Gly
 435 440 445
 Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp Ala Leu Ala
 450 455 460
 Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly Arg Tyr Pro
 465 470 475 480
 Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp Ser Lys Pro
 485 490 495
 Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr Gly Lys Gly
 500 505 510
 Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro Val Ala Glu
 515 520 525
 Lys Ile Pro
 530

<210> 10
 <211> 1696
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: cDNA coding for the
 carboxy terminal sequence for the GBSSI of
 Chlamydomonas reinhardtii

<220>
 <221> CDS
 <222> (3)..(716)

<400> 10
 at tcg gca cga gtg cac gcc atg cac tac ggt acc gtg ccc gtg gta 47
 Ser Ala Arg Val His Ala Met His Tyr Gly Thr Val Pro Val Val
 1 5 10 15
 gcc tcc acc ggc ggc ctg gtc gac acc gtc aag gag ggc gtc acc ggc 95
 Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly
 20 25 30
 ttc cac atg ggc gcc ctg aac ccc gac aag ctg gac gag gct gac gcc 143
 Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala
 35 40 45
 gac gcc ctg gcc gcc acc gtg cgc cgt gcc agc gag gtg ttt gcg ggc 191
 Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly
 50 55 60
 ggc cgc tac ccc gag atg gtg gcc aac tgc atc agc cag gac ctg tcc 239
 Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser
 65 70 75

tgg tcc aag ccc gcc cag aag tgg gag ggc ctg ctg gag gag gtg gtg	287
Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val	
80 85 90 95	
tac ggc aag ggc ggc gtg gcc acc gcc aag aag gag gag atc aag gtg	335
Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val	
100 105 110	
ccc gtt gcc gag aag atc ccc ggc gac ctg ccc gcc gtg tcc tac gcc	383
Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro Ala Val Ser Tyr Ala	
115 120 125	
ccc aac acc ctg aag ccc gtg tcc gcc tcc gtg gag ggc aac ggc gcc	431
Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val Glu Gly Asn Gly Ala	
130 135 140	
gcc gcg ccc aag gtc ggc acc acc gcc ccc gcc atg ggc gcg tgg cgc	479
Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala Met Gly Ala Trp Arg	
145 150 155	
gcg acc acc ccc tcg ggc ccc tcg ccc gcc gcc gcc acc ccc aag gtg	527
Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala Ala Thr Pro Lys Val	
160 165 170 175	
acc acc tac aag ccc gcc ctg ccc gcc acc gcc aag ccc aag acc gct	575
Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala Lys Pro Lys Thr Ala	
180 185 190	
ggc ctc aag ctg gcc ggt gag gcc tcc acc acc tcg acc tcg gag aac	623
Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr Ser Thr Ser Glu Asn	
195 200 205	
ggc gct gcc tcc aac ggc aac ggc aac ggt gcc tcg gcc tcc aag acc	671
Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala Ser Ala Ser Lys Thr	
210 215 220	
tcg gct gcc aag ccc ctg gtc tcc gcc gcc acc cgc aag tcc gcc	716
Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr Arg Lys Ser Ala	
225 230 235	
taaagcggca gtagccgcag aggcggcgac agcatgagcg gctcgaccaa agctgtggca	776
ggaacggctg tagcagcggc aggcggccgc caccggcgag gagcaggctt gcggcagcga	836
gggcgatgag cttagcggcc gtgagcatgg caggcggaaa cgtgtgtact gaaatgtggt	896
gcatgagagt gtcgtgctgt aatgaagtcg gttttgagcg accggagaaa cgccggtttg	956
gttttgtagt gcagggcctg tggtttcggt tttgcccaag tccaaaagaa gagtaacgaa	1016
actgtagcag tagcagagca cttgcgcggc gcggcgacca cgccggcccg tgcgagcct	1076
gtcctgccct cagccttggt attcggcggc aagagggcgg gtctgtacac tccatccatt	1136
ccaggatttt tgcaggctgc ctgagagttt gccattttgt gggacgtgag cggcgggacg	1196
gccgcgccgg gctctcctac cgcctccggc aacggagaag tgggagggcg ttagccccg	1256

```

tgacccccca atgtagagga tgggatacat aagagcgtgt ggaatggtgg taaaagagga 1316
ggggcctggg tcgcccctcg atgggttttgt tgaggtgcag acggcacctg cggcgtcaaa 1376
ggccctcgca agggccgggt gccttgggct cttttttggt gcccgtcgat gatgagagat 1436
tggccagcgg ttttttgagg ctggctcgaa gcgaggggtt gtggaagtgg agcgaggagg 1496
gttgagaaaa gaggcggaca tgcttgactg gaggtacaca aagtggagcg tgcgacggca 1556
cggaggcatt ggcggactat tgaccagta gtgtggaaag tagttggacc tgaattcttt 1616
gagagtaccg cgcattaatc cgtgagagag taacaaagat ggcacctgaa aaaaaaaaaa 1676
aaaaaaaaaa aaaaaaaaaa 1696

```

<210> 11

<211> 1696

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: cDNA coding for the
carboxy terminal sequence for the GBSSI of *Chlamydomonas*
reinhardtii

<400> 11

```

tttttttttt tttttttttt tttttttttt ttcaggtgcc atctttgtta ctctctcacg 60
gattaatgcg cgtactctc aaagaattca ggtccaacta ctttccacac tactgggtca 120
atagtccgcc aatgcctccg tgccgtcgca cgctccactt tgtgtacctc cagtcaagca 180
tgtccgcctc tttctccaac cctcctcgct ccacttccac aaacctcgc ttcgagccag 240
cctcaaaaaa ccgctggcca atctctcatc atcgacgggc accaaaaatg agccaaggc 300
acctgggctc tgcgagggcc tttgacgccg acggtgccgt ctgcacctca acaaaacct 360
cgagggcgca cccaggcccc tcctctttta ccaccattcc acacgctctt atgtatcca 420
tcctctacat tggggggtca ccgggctaca gcgcctccca cttctccgtt gccggaggcg 480
gtaggagagc ccggcgcggc cgtcccgcgc ctcacgtccc acaaaatggc aaactctcag 540
gcagcctgca aaaatcctgg aatggatgga gtgtacagac ccgccctctt gccgccgaat 600
cacaaggctg agggcaggac aggtgcgca cgggccggcg tggtcgccgc gccgcgcaag 660
tgctctgcta ctgctacagt ttcgttactc ttcttttgga cttgggcaaa accgaaacca 720
caggccctgc actacaaaac caaacggcg tttctccggt ctgcgaaaac cgacttcatt 780
acagcacgac actctcatgc accacatttc agtacacacg tttccgctg ccatgctcac 840
ggccgctaag ctcatcgccc tcgctgccgc aagcctgctc ctgcgccgtg gcggccgcct 900
gccgctgcta cagccgttcc tgccacagct ttggtcgagc cgctcatgct gtcgccgcct 960
ctgcggtac tgccgcttta ggcggacttg cgggtggcgg cggagaccag gggcttggca 1020
gccgaggtct tggaggccga ggcaccgtg ccgttgccgt tggaggcagc gccgttctcc 1080
gaggtcgagg tgggtgaggc ctcaccggcc agcttgaggc cagcggctct gggcttggcg 1140
gtggcgggca gggcgggctt gtaggtgttc acctggggg tggcgggcgg gggcgagggg 1200
cccgagggg tggtcgcgcg ccacgcgccc atggcggggg cgggtggtgc gacctgggc 1260
gcggcgggcg cgttgccctc cacggaggcg gacacgggct tcagggtgtt gggggcgtag 1320
gacacggcgg gcaggtcgcc ggggatcttc tcggcaacgg gcacctgat ctctctcttc 1380
ttggcggtgg ccacgccgcc cttgccgtac accacctcct ccagcaggcc ctcccacttc 1440
tgggcgggct tggaccagga caggtcctgg ctgatgcagt tggccaccat ctcggggtag 1500
cggccgcccc caaacacctc gctggcacgg cgcacggtgg cggccagggc gtcggcgta 1560
gcctcgcca gcttgctcgg gttagggcg cccatgtgga agccggtgac gccctccttg 1620
acggtgtcga ccaggccgcc ggtggaggct accacgggca cggtagcgta gtgcatggcg 1680

```


<210> 12
 <211> 238
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Carboxy terminal
 peptide sequence of the GBSSI of Chlamydomonas
 reinhardtii

<400> 12
 Ser Ala Arg Val His Ala Met His Tyr Gly Thr Val Pro Val Val Ala
 1 5 10 15
 Ser Thr Gly Gly Leu Val Asp Thr Val Lys Glu Gly Val Thr Gly Phe
 20 25 30
 His Met Gly Ala Leu Asn Pro Asp Lys Leu Asp Glu Ala Asp Ala Asp
 35 40 45
 Ala Leu Ala Ala Thr Val Arg Arg Ala Ser Glu Val Phe Ala Gly Gly
 50 55 60
 Arg Tyr Pro Glu Met Val Ala Asn Cys Ile Ser Gln Asp Leu Ser Trp
 65 70 75 80
 Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu Leu Glu Glu Val Val Tyr
 85 90 95
 Gly Lys Gly Gly Val Ala Thr Ala Lys Lys Glu Glu Ile Lys Val Pro
 100 105 110
 Val Ala Glu Lys Ile Pro Gly Asp Leu Pro Ala Val Ser Tyr Ala Pro
 115 120 125
 Asn Thr Leu Lys Pro Val Ser Ala Ser Val Glu Gly Asn Gly Ala Ala
 130 135 140
 Ala Pro Lys Val Gly Thr Thr Ala Pro Ala Met Gly Ala Trp Arg Ala
 145 150 155 160
 Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala Ala Thr Pro Lys Val Thr
 165 170 175
 Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala Lys Pro Lys Thr Ala Gly
 180 185 190
 Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr Ser Thr Ser Glu Asn Gly
 195 200 205
 Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala Ser Ala Ser Lys Thr Ser
 210 215 220

Ala Ala Lys Pro Leu Val Ser Ala Ala Thr Arg Lys Ser Ala
 225 230 235

<210> 13
 <211> 23
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Amino terminal
 fragment of GBSSI of Chlamydomonas reinhardtii

<400> 13
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Gly Gly Lys
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val
 20

<210> 14
 <211> 23
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Amino terminal
 fragment of GBSSI of Chlamydomonas reinhardtii

<400> 14
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val
 20

<210> 15
 <211> 3117
 <212> DNA
 <213> Chlamydomonas reinhardtii

<400> 15
 tttttttttt tttttttttt tttttttttt tcaggtgcc a tctttgttac tctctcacgg 60
 attaatgcgc ggtactctca aagaattcag gtccaactac ttccacact actgggtcaa 120
 tagtccgcc aatgcctccgt gccgtcgac gctccacttt gtgtacctcc agtcaagcat 180
 gtccgcctct ttctccaacc ctctcgtc cacttccaca aacctcgtc tcgagccagc 240
 ctcaaaaaac cgctggccaa tctctcatca tcgacgggca ccaaaaatga gccaaggca 300
 cccgggcctt gcgagggcct ttgacgccga cgggtgccgc tgcacctcaa caaaaccatc 360
 gaggggcgac ccaggccctt cctcttttac caccattcca cacgctctta tgtatcccat 420
 cctctacatt ggggggtcac cgggctacag cgctccac ttctccgttg ccggaggcgg 480
 taggagagcc cggcgccgcc gtcccgccgc tcacgtcca caaaatggca aactctcagg 540
 cagcctgcaa aaatcctgga atggatggag tgtacagacc cgccctcttg ccgccgaatc 600
 acaaggctga gggcaggaca ggctgcgcac gggccggcgt ggtcgccgcg ccgcgcaagt 660
 gctctgctac tgctacagtt tcgttactct tcttttgac ttgggcaaaa ccgaaaccac 720
 aggccctgca ctacaaaacc aaaccggcgt ttctccggtc tcgcaaaacc gacttcatta 780

cagcacgaca	ctctcatgca	ccacatttca	gtacacacgt	ttccgcctgc	catgctcacg	840
gccgctaagc	tcacgcgcct	cgctgccgca	agcctgctcc	tcgccggtgg	cggccgcctg	900
ccgctgctac	agccgttcct	gccacagctt	tggtcgagcc	gctcatgctg	tcgccgcctc	960
tgcggctact	gccgctttag	gcggacttgc	gggtggcggc	ggagaccagg	ggcttggcag	1020
ccgaggtcct	ggaggccgag	gcaccgttgc	cggtgccggt	ggaggcagcg	ccgttctccg	1080
aggtcgaggt	ggtggaggcc	tcaccggcca	gcttgaggcc	agcggctctt	ggcttggcgg	1140
tggcgggcag	ggcgggcttg	taggtggtca	ccttgggggt	ggcggcggcg	ggcgaggggc	1200
ccgagggggg	ggtcgcgcg	cacgcgccca	tggcgggggc	ggtggtgccg	accttgggcg	1260
cggcggcgcg	ggtgcctcc	acggaggcgg	acacgggctt	caggggtgtg	ggggcgtagg	1320
acacggcggg	caggtcgccg	gggatcttct	cggcaacggg	caccttgatc	tcctccttct	1380
tggcgggtgg	cacgccgccc	ttgccgtaca	ccacctcctc	cagcaggccc	tcctccttct	1440
gggcgggctt	ggaccaggac	aggtcctggc	tgatgcagtt	ggccaccatc	tcggggtagc	1500
ggccgcccgc	aaacacctcg	ctggcacggc	gcacggtggc	ggccaggcg	tcggcgtagc	1560
cctcgctccag	cttgctgggg	ttcagggcgc	ccatgtggaa	gccggtgacg	ccctccttga	1620
cgggtgtcgac	caggccgcgc	gtggaggcta	ccacgggcac	ggtaccgtag	tgcatggcgt	1680
gcagctggat	caggccgcag	ggctcgaagc	gcgagggcac	cagcatgaag	tcggcgcccg	1740
cgggtagcat	gtgcgccagg	ggcgccgaga	acttgaccac	gcccttggcg	cggcccttgt	1800
acttgggtgcc	gatggcgctt	accagcttct	cgtaggcggc	cttgccggtg	cccaggatgg	1860
cgatctgcac	cttgggggtg	gccaggatct	tgggcagggc	ggccaggatg	atgtccacac	1920
ccttctgtct	ctccaggcgc	ccgatgaagg	cgaacagggg	ggcgggtggg	tcacacaggca	1980
ggcccagctc	ggcctgcagg	gcctccttgg	cggcggcctt	gccggcgtag	acgctgttct	2040
ggttcgaagg	cgcagacagg	aacttgtcgg	tcttgggggt	ccactcctca	atgtccatgc	2100
cggtcacaat	gccttcaatg	cccttggcgc	ggatgcagg	gtccagctcc	acaccgccgg	2160
cggcatcggc	agcgatctcg	gtcgcgtagt	tgggcgacac	agtcaccagc	ttgtcggcgg	2220
cgataatgcc	acccttcagc	cagttgatct	tcttgtaggt	ctttcccgtc	agcgggggct	2280
tctcgtcctc	ctccatgggg	gtggcctcag	tgtaaacctt	ggcatagccg	tcggagaagg	2340
ccagcttgct	aaaggcggcc	tggggcagct	tcgtgtcctt	gaaagcctcc	tcacacatgc	2400
ggccctggaa	ggcgatggtg	tggatagcca	gcaccgactt	ggccttgggt	aactggccct	2460
tgggctggta	ctcgtccttc	agcaggacgg	gcaccagggc	ggagtgccag	tcgttggcca	2520
cgaagacgca	gtcctcgccg	gggcccgaag	gcagcacgcg	ggcagcctca	atagcggcct	2580
tgcagaacag	ggcgaagcgc	ttgtggttgt	ccaggtagtc	agcgccggag	cgggggcccgt	2640
acagcttgga	gccggtcctt	ccccagacct	tggccaggaa	ccaggggtgg	tcaatccaca	2700
cgcggtgcac	gcccttcttg	atggagtgga	agtagcgac	cttctcgccc	atgatgtcca	2760
cgaccaccga	ggtgtcccag	gcgtcagcgt	actggtcgta	gcgaggggca	atggtcatga	2820
cgcggtggcc	gcgcttgacc	agctcaatag	gcaggccacc	agtcacatcg	cccaggccgc	2880
ccgtcttgga	ccaaggggcg	acctcagcag	caaccatcac	gatgtccagc	gcgcaagtgg	2940
caccagtggc	accagtaaca	gccgagcgcg	aggtggactt	gcgtgcggag	ccacgagcaa	3000
gctcagcgag	cagctggttc	gcggtcttct	tgacaccgaa	cgacgcggca	ttgatgacga	3060
taggacgcgc	gctgctgggg	cggctggtag	aggcaacagc	catacgctcc	gcctggc	3117